DOCUMENT RESUME

ED 150 216 UD 017 838

AUTHOR Richardson, Debra Jeanne
TITLE Internal Consistency of the Test for Auditory Comprehension of Language in Headstart Children.
PUB DATE 76
NOTE 35p.

EDRS PRICE MF-$0.83 HC-$2.06 Plus Postage.
DESCRIPTORS Black Youth; *Diagnostic Tests; *Language Ability; *Language Skills; Low Income Groups; *Preschool Children; Preschool Education; Testing Problems; *Test Reliability
IDENTIFIERS Arkansas (Pulaski County); *Test for Auditory Comprehension of Language

ABSTRACT

The internal consistency of the test for Auditory Comprehension of Language (TACL), as it is used in the diagnostic assessment of the language comprehension of black, low socioeconomic preschool children, was investigated in this study. 30 three year old and 30 four year old black Headstart children from Pulaski County, Arkansas were given the TACL after performing satisfactorily on a language screening test. Each child was given the test at his or her Headstart center according to published instructions. Raw scores were collected and analyzed for the test according to the alpha coefficient methods. Results indicated that the test has only a moderate level of internal consistency for use with low socioeconomic, black preschool children, a level less than desirable for satisfactory language assessment. There is no explanation for this level of internal consistency, and some of the factors that may have contributed to the results include the possible sources of error which contribute to lack of reliability in tests: instrument error, subject error, and administrative error. Some possible sources of subject error include variation within subjects in guessing, state of motivation, and language differences. The content of the test as a source of error indicates the need for an item analysis of the TACL. This study also emphasizes the need for assessment devices which measure language performance according to normative data established for the group and data reflecting internal consistency within the structure of the tests. (Author/AM)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED150216

INTERNAL CONSISTENCY OF THE TEST FOR AUDITORY COMPREHENSION OF LANGUAGE IN HEADSTART CHILDREN

A research project submitted in partial
fulfillment of the requirements for
the degree of Master of Sciences

by

Debra Jeanne Richardson, B.A.
University of Arkansas, 1975

UP017838

1976
The University of Arkansas for
Medical Sciences and the
University of Arkansas at Little Rock

U S DEPARTMENT OF HEALTH,
EDUCATION & WELFARE
NATIONAL INSTITUTE OF
EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

Debra Richardson

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC) AND USERS OF THE ERIC SYSTEM."

ABSTRACT

This research examined the internal consistency of the Test for Auditory Comprehension of Language (TACL), (Carrow, 1973) as it is utilized with Headstart children in diagnosing language problems. The study involved administering the TACL to 30 3-year-old and 30 4-year-old black children selected randomly from Headstart centers.

Administrative controls were implemented to minimize examiner and administration error. These controls included a schedule for reinforcement, consistent stimulus delay time, and administration by one examiner.

To obtain internal consistency measures for the TACL, the test was subdivided into equal halves and correlated for an estimate of test reliability. The alpha coefficient, a statistical method for determining objective consistency measures, was employed.

The results of this study indicated that the TACL has a less than desirable level of internal consistency for use with Headstart children, as shown in correlation coefficients of .59 and .66 as well as narrow ranges of scores and standard deviations. This data differed from the normative data established for the same age groups. These results indicate the need for further research into the internal structure of tests such as the TACL. On a broader perspective, the results of this research show that there is a need for further construction of adequate diagnostic devices for measuring language performance of Headstart children.

TABLE OF CONTENTS

	Page
ABSTRACT	iii
LIST OF TABLES	vii
CHAPTER	
I. INTRODUCTION	1
II. METHOD	10
III. RESULTS	14
IV. DISCUSSION	19
V. SUMMARY	25
APPENDIX	27
REFERENCES	28

LIST OF TABLES

TABLE		Page
I.	DESCRIPTION OF SUBJECTS	11
II.	MEANS AND STANDARD DEVIATIONS FOR HALF AND TOTAL TEST ...	15
III.	NORMATIVE AND PRESENT STUDY MEANS AND STANDARD DEVIATIONS FOR HALF AND TOTAL TEST (TACL)	16
IV.	CORRELATION COEFFICIENTS (ALPHAS) FOR PRESENT STUDY ...	17

INTROLUCTION

Purpose

With the increasing interest in the development of preschool language intervention programs designed for low socioeconomic, black children in the last decade, there have also been serious attempts made to establish valid language assessment devices appropriate to that population. Through this concern a number of issues have been introduced, among them the reliability of standardized language tests for use with low socioeconomic, black children. The purpose of this research was to examine such an issue, utilizing the Test for Auditory Comprehension of Language (Carrow, 1973) with Headstart children.

The Test for Auditory Comprehension of Language is among the tests currently used to assess oral language comprehension in children. It was developed by Dr. Elizabeth Carrow to describe the order in which a child is expected to master a particular grammatical form. The test consists of 101 plates, each containing black and white line drawings. One of the three drawings is the referent for the linguistic form being tested. The alternative drawings represent the referent for a contrasting linguistic form or a single contrast plus a decoy. The structure and function of nouns, verbs, adjectives, adverbs, and prepositions are assessed as well as morphological constructions, grammatical categories involving case, number, tense, status, voice and mood, and syntactic structures of predication,

complementation, modification, and coordination. The subject is asked to point to the picture illustrating the stimulus word given by the examiner.

The TACL has been found to successfully distinguish between those individuals who have disorders of language with those who do not using the normative data (Carrow and Lynch, 1975). The test has published normative data obtained from a selected sample of 200 middle socioeconomic status Black, Anglo, and Mexican-American children between the ages of three to six years.

This research employed the TACL with a group of black, low socio-economic children to investigate the reliability of the TACL as a measure of language comprehension for Headstart subjects. Reliability was tested by a correlation of test items for internal consistency. This study, in addition, contributes data concerning variability of scores, and discusses possible sources of measurement error.

Review of the Literature

Because measurement of error is an important issue in the use of any testing instrument, investigations of reliability are routinely made when new devices are developed. Most studies agree that the rationale for test reliability is a basic concern for the accuracy with which a score represents an individual, and therefore is repeatable and free from error.

Those studies regarding test measurement have agreed that rarely is measurement free from error (Thorndike, 1969; Kerlinger, 1964). In addition, there are many sources of error. They can be categorized into instrument error, situation error, and performance or administrative error.

Struening and Guttentag (1975) have pointed out, for example, that for multiple-choice tests such as the TACL, guessing is often a source of measurement error that may cause variation in performance from item to item, and which may tend to lower the test reliability.

Internal consistency represents an estimate of reliability based on the average correlation among the items within a test. The basic formula for determining reliability based on internal consistency is the coefficient alpha (Kerlinger, 1964). This index of reliability provides a good estimate of the effectiveness of an instrument since the major source of measurement error is sampling of content within the instrument. Kuder-Richardson Formula 20 is the most acceptable version of coefficient alpha for the establishment of a reliability correlation (Kerlinger, 1964).

To estimate error of measurement in a single testing, most researchers agree that the most widely used procedure is sub-division of the test into equal halves (Campbell, 1963; Wick, 1973; Kerlinger, 1964). By correlating the two half-length tests one is able to compute an unbiased estimate of total test reliability.

The Test for Auditory Comprehension of Language as a Diagnostic Method

The Test for Auditory Comprehension of Language (TACL) is currently employed in many clinical and educational preschool settings as a method for diagnosing language comprehension problems. Among the settings are Headstart centers. In reviewing the literature related to the reliability of the TACL, almost no data have been reported that consider the low socioeconomic Headstart population. Norms for these children were not developed in the construction of the test; yet the TACL is widely used

within Headstart centers to assess language performance.

One study (Jones, 1975) has investigated oral language comprehension of black and white middle and lower class preschool children in standard English and black dialect using the TACL as a method of investigation. A cross sectional sample of 80 preschool children randomly selected for socioeconomic status were chosen and equally divided between the 4 and 5-year-old range. The study translated the TACL into black dialect, and found that there were significant differences in the levels of oral language comprehension in both standard and nonstandard English when subjects were grouped by socioeconomic status. This indicated that social class was a major determining factor in the young child's ability to comprehend language structures, a fact that has been repeatedly proven in the studies related to low socioeconomic children's language comprehension (Baratz, 1969; Gerber and Hertel, 1969; Osser et al, 1969; Miner and Shriner, 1968). The study also regarded age as a significant factor in the oral language comprehension of most subjects in both standard and nonstandard dialects.

While the Jones study is valuable for the indications made by its results, there are several issues not addressed. 3-year-olds were not included in the study of preschool subjects. Questions concerning the individual test items and their contribution to total test performance were not considered. Variation in administration and subject variation were not considered. The study investigated primarily differences in the language comprehension between standard and nonstandard English with the TACL as an adjunct to investigation.

Other research related to the TACL has considered it as a method

of investigation, rather than investigation of the TACL itself. Braun (1975) studied the possible correlation of the syntax portion of the TACL to the receptive subtest of the Northwestern Syntax Screening Test (NSST). The study did not demonstrate a correlation, but considering that the NSST has been shown to lack internal consistency (Ratusnik, 1975), and that the TACL presents only limited information on syntax (12 of its 101 items), the low correlation may reflect different language skills for the NSST than are required for the TACL (Braun, 1975), or a weakness in the structure of one or both tests.

Another study involving the TACL as a research method (Carrow and Lynch, 1975) investigated the performances of three linguistically deviant groups of 5-year-olds on the TACL and the Peabody Picture Vocabulary Test. The results showed that the two tests measure two different skills of language comprehension, and therefore the children performed differently on each test.

With the exception of a study by Musselwhite (1975), the research related to the TACL has dealt with measures external to the test itself by either comparing the test to another assessment device or using it to describe specific language groups. Musselwhite has presented normative data on the Test for Auditory Comprehension of Language for lower socioeconomic status children. By administering the TACL to 299 preschool Headstart children and interpreting raw scores according to means and standard deviations, the author showed statistically that scores for low socioeconomic status subjects varied considerably from the middle socioeconomic status norms. Low socioeconomic status children were found to score on the average one or more standard deviations below

the mean of their middle socioeconomic counterparts. This study has provided much useful data regarding the TACL. In addition to establishing norms for low socioeconomic children, the study questioned the validity of using standardized assessment devices such as the TACL with the low socioeconomic preschool population without the use of additional controls to minimize administrative error. Periodic social reinforcement was delivered for appropriate attending behavior during administration of the TACL (Musselwhite, 1975). This issue had not previously been considered in dealing with the TACL and low socioeconomic children, despite the emphasis on the importance of positive reinforcement in dealing with low socioeconomic children (Dickie et. al., 1972; Ratusnik etl al., 1975; Cazden, 1966; Baratz, 1969).

In summary, the literature concerning the TACL has primarily been concerned with the test as a method of research or diagnosis, and not with the structure of the test itself. One study, (Musselwhite, 1975) has dealt with issues internal to the TACL as it is used with low socioeconomic preschool children, and has questioned how adequately language comprehension of these children is measured when standardized assessment devices which do not provide corresponding normative data are used.

Language Comprehension Ability in Low Socioeconomic Black Children

Related to the concerns for further needed research in the use of standardized assessment devices such as the TACL with low socioeconomic preschool children is the issue of low socioeconomic children's language ability. For years there has been a controversy regarding whether the language ability in these children is deficient or simply different.

Somervill (1974) pointed out that not enough is known about the language of low socioeconomic, black preschool children to make any judgements about it. "Before we can remediate a language deficit, we must know what the language is and what it conveys." (Somervill, 1974). In addition, the literature related to low socioeconomic children's language does not always differentiate between what is meaning and what is structure. The aspect of language which has been described by educators and psychologists as deficient is usually meaning. The aspect of language which has been defended by linguists as being different is structure. Until this confusion is recognized and dealt with conceptually, there is no hope for a resolution of these two points of view (Somervill, 1974).

Accordingly, there is ambiguity in the literature regarding whether or not the language difficulty exhibited by low socioeconomic children interferes with their ability to comprehend standard English as it is presented on many standardized assessment devices. Some authors have pointed out that the interference with language comprehension is only an assumption (Baratz, 1969; Somervill, 1974). Others have indicated that it is the particular speech situation itself, the methodology of the language research, or the interpretation of data which have resulted in the conclusion that the language of low socioeconomic children interferes with comprehension (Osser et. al., 1969; Ervin-Tripp, 1971; Cazden, 1971; Dickie and Bagur, 1972).

The few studies which have dealt with the language comprehension of low socioeconomic children have done so on a comparative basis, using the white middle class usually. Osser, Wang, and Zaid (1969) concluded from a study testing for imitation and comprehension that black children's control over

selected common syntactic structures in standard English is markedly inferior to that of white subjects, even when dialect variations were considered. This study has been one of the few to assess linguistic differences of low socioeconomic, black children in tasks other than production.

Ratusnik and Koenigskecht (1974) examined certain clinical variables and their influence on black preschoolers nonstandard phonological and grammatical performance. The authors found that the lower socioeconomic group evidenced many more grammatical forms which were nonstandard than did the middle socioeconomic group, and that while not statistically significant, boys employed substantially more nonstandard constructions than girls. This, too, was most noticeable in the lower socioeconomic group.

Language differences in the low socioeconomic preschool group have been examined by several authors: Stern and Gupta (1970) examined echoic responding as a function of type of speech modeled, and found no difference between performance of black children's responding on "dialect" and standard versions of the assessment. Osser, Wang, and Zaid (1969) stressed the need for looking at both comprehension and production at the same time in order to determine the child's knowledge of certain linguistic forms. In some instances the authors found that children could repeat forms that they did not appear to comprehend and in other instances they could comprehend forms that they failed to reproduce on the imitation task.

Regardless of the deficit-difference issue, studies by Baratz (1971) and Ervin-Tripp (1970) have demonstrated that research on low socioeconomic

children's comprehension is greatly needed. They have indicated that too often comprehension is not assessed validly because a sentence repetition or auditory discrimination task is used that assess production and phonology.

It is important to note that the literature concerning low socioeconomic children's comprehension of language is insufficient to provide any generalizations about the language comprehension ability of that population. The studies which are available are somewhat limited in methodology by the use of tasks which are more related to production or some other ability than comprehension, and there have been many issues not addressed in the literature related to this area.

The issues considered crucial and therefore investigated in this study of the internal consistency of the Test for Auditory Comprehension of Lanugage were those of test content and structure as it affects the performance of low socioeconomic, black preschoolers and the internal consistency of a standardized language test such as the TACL as it may vary language performance.

The purpose of this research was to investigate the reliability of the TACL testing for internal consistency with a selected group of black, low socioeconomic subjects from the Headstart population. The TACL was administered to 60 Headstart children with no speech, language, or hearing problem to determine how consistently the test measures language comprehension in the low socioeconomic, black preschool group.

Chapter II

METHOD

Subjects

To examine the question whether or not the TACL is internally consistent when utilized with Headstart children, 60 black, low socioeconomic children were chosen from various Headstart centers of Pulaski County, Arkansas. Two equal groups of 30 children were chosen, with 30 children in the 3-year-old group and 30 children in the 4-year-old group. Subjects were screened, and no speech, language, or hearing impairment was identified. Language was assessed by the Utah Test of Language Development (Mecham, 1967), a diagnostic language test designed to assess receptive and expressive language ability. Males and females were included in the study. A description of the subjects is provided in Table I. Group A subjects were 3-year-old children, and Group B subjects were 4-year-old children.

Procedure

Each subject was administered the Test for Auditory Comprehension of Language (TACL) on one occasion at his or her respective Headstart center. The test required the subject to point to 101 line drawing stimuli appropriately as a stimulus word was given by the examiner. Each test was given in approximately 15 minutes, according to the procedure outlined in the TACL test manual.

TABLE I

DESCRIPTION OF SUBJECTS

AGE	MALE	FEMALE	TOTAL
3 years (A)	11	19	30
4 years (B)	14	16	30

Additional administrative controls were systematically included to reinforce attending behavior and to control length of time that stimuli were presented. Beginning with item #1, every tenth consecutive item was socially reinforced, using verbal praise in the form of "you're doing a good job." This procedure was continued for the entire test. Consistent three second delay time was utilized for each subject in the presentation of each stimulus. Along with social reinforcement, an emphasis was placed on controlling for random responses by asking the child to remember to "look at all of the pictures on the page."

Each subject was coded for age and sex by numeral and letter to represent the order in which the child was tested. This information, along with test data, was recorded on a revised record form of the Test for Auditory Comprehension of Language. (See Appendix) Items marked by an asterisk (*) indicate those items where social reinforcement was delivered. This score form was adapted for ease of statistical treatment. Preparation was made for dividing the test into odd and even test items, and for having order of presentation consecutive.

Each test was individually scored after all tests were administered. Incorrect item numbers were slashed (/), and the subjects' responses were recorded by number in the corresponding blank beside the correct item. The incorrect responses were totaled and subtracted to obtain raw scores.

Analysis and Interpretation of Data

Using the alpha coefficient, the general form for computing internal consistency measures (Kerlinger, 1964), the following formula was employed:

$$\alpha = 2\left(\frac{1-s^2_{\text{odd}} + s^2_{\text{even}}}{s^2_{\text{total}}}\right)$$

To compute internal consistency measures with this formula, it was necessary to first divide the TACL into equal halves, then determine half-test raw scores, total test raw scores, means, and standard deviations for each subject respectively. Using this data, the alpha coefficient formula enabled the computation of resulting correlation coefficients for each age group (Table IV).

Chapter III

RESULTS

The internal consistency of the Test for Auditory Comprehension of Language was estimated by the utilization of the alpha coefficient (Kerlinger, 1964), the general form for estimating internal consistency. The formula is read as follows:

$$\alpha = 2\left(\frac{s^2 \text{ odd} + s^2 \text{ even}}{s^2 \text{ total}}\right)$$

where s^2 = variance.

In utilizing this formula, it was necessary to subdivide the TACL into half-tests by placing the odd numbered items in one group and the even numbered items in another group. The resulting raw scores were used to obtain a correlation for the total test.

Table II presents the means and standard deviations for each half-test of the TACL and the total test respectively as obtained by this research. As it indicates, both Group A (3-year-olds) and Group B (4-year-olds) scored higher on half-test I than on half-test II, and the average score for Group A for the total test was 52.

Table III compares the means and standard deviations for Group A and Group B in the present study with those of the normative study (Carrow, 1973). As indicated, both means and standard deviations were reported higher in the normative data of the TACL than in the present study. The differences for means was greater than nine points for Group A and 15 points for Group B

TABLE II

MEANS AND STANDARD DEVIATIONS FOR HALF AND TOTAL TEST (TACL)

GROUP	HALF-TEST 1 $\bar{X}$	HALF-TEST 1 S.D.	HALF-TEST 2 $\bar{X}$	HALF-TEST 2 S.D.	TOTAL TEST $\bar{X}$	TOTAL TEST S.D.
A 3-year	22.50	3.49	22.40	3.02	45.26	5.50
B 4-year	26.93	5.42	25.96	4.29	52.90	8.46

TABLE III

MEANS AND STANDARD DEVIATIONS FOR TOTAL TEST (TACL)

	GROUP	NORMATIVE $\bar{X}$	NORMATIVE S.D.	MUSSELWHITE $\bar{X}$	MUSSELWHITE S.D.	PRESENT STUDY $\bar{X}$	PRESENT STUDY S.D.
A	3-0 to 3-5	54.30	8.76	42.92	5.47	45.26	5.50
	3-6 to 3-11	62.10	11.69	47.82	9.21		
B	4-0 to 4-5	69.37	11.05	55.69	12.00	52.90	8.46
	4-6 to 4-11	69.64	9.70	57.67	9.07		

TABLE IV

CORRELATION COEFFICIENTS (ALPHAS)

GROUP A	(3-year-olds)	= .591441
GROUP B	(4-year-olds)	= .662602

Standard deviations were smaller for subjects in this study than for the normal group. In addition, correlations in this study by age group are reported lower than in the TACL normative data.

Table IV presents correlation coefficients for each age group in this study. A correlation of .59 was reported in this study for the 3-year-olds (Group A), and a correlation of .56 was reported for 4-year-olds (Group B). These findings vary from the .77 correlation reported in the TACL normative data study.

Table III shows how data from this research compare to data reported by Musselwhite (1975) for low socioeconomic preschool children. She reported that the means and standard deviations for her subjects fell below the norms of their middle socioeconomic counterparts by at least one standard deviation. Some data from the present study match that collected by Musselwhite. Differences in means and standard deviations are noted between the two studies, however. One major difference is the reduced range of scores for both groups of subjects in this study as contrasted to the Musselwhite study. Group B subjects also had a lower mean score than she found in the 4-year-old subjects in North Carolina.

The results of the statistical analysis for internal consistency of the TACL for selected Headstart subjects in Pulaski County, Arkansas was a correlation of .59 for 3-year-old subjects and .66 for 4-year-old subjects.

Chapter IV

DISCUSSION

This study examined the internal consistency of the Test for Auditory Comprehension of Language (TACL) when utilized with black, low socioeconomic preschool children. The results of this research indicate that the TACL has only moderate reliability, a level of internal consistency less than desirable as it is used with the low socioeconomic, black population.

Since there is, as yet, no clear explanation for this moderate level of internal consistency in the TACL, discussion of the factors which may have contributed to the results of this study includes the possible sources of error which contribute to lack of reliability in tests: instrument error, subject error, and administrative error.

Test Administration

In an attempt to control for the effects of variation in administration of the test, this study pre-determined procedures for consistent test administration. The TACL was administered to each subject in the study according to the instructions outlined in the test manual with additional controls for stimulus presentation and reinforcement. In an effort to assure the attention of the subject, positive reinforcement in the form of verbal praise was presented consistently to every subject throughout the testing in the same

manner and at the same intervals. A single examiner presented the language test.

Musselwhite (1975) examined the use of social reinforcement in the administration of the TACL to low socioeconomic, black preschoolers. Positive reinforcement was delivered randomly rather than systematically, however. This may have contributed to the results obtained in her study, if one subject received more or less reinforcement than another subject. Variation in results may also have been present in the Musselwhite study by the use of several examiners, rather than one, as the present study did to further control for administrative error.

Other administrative controls included an attempt to regulate random responses in subjects through the use of verbal clues that were delivered in a systematic fashion, on the same items where positive reinforcement was delivered. The verbal clues given by the examiner consisted of a reminder to "look at all of the pictures on the page." To eliminate possible examiner bias, no individual test forms were scored until all 60 tests were administered.

One factor within the category of administrative error which may have contributed to variation in performance in this study is the race of examiner, which differs from the subjects' in this study. There is general agreement in the literature (Adler, 1971) that variation in subject performance may occur as a result of this factor. Despite this effect, it is believed that clinical experience in the form of two semesters clinical practicum work with these children in various Headstart centers results in some degree of knowledge on the part of

the experimenter in establishing a relationship of rapport and confidence with the children on an individual basis. This experience, it is believed, would modify the factor.

In view of these controls, it is doubtful that administration of the TACL served as a major source of error contributing to the moderate level of internal consistency obtained for use with black, low socioeconomic preschool children.

Subject Performance

There are numerous possibilities that exist as sources of error within the performance of subjects, due to the obvious individuality of human subjects. One course within subject variability that this study has indicated relevant for consideration is that of guessing. This factor, as well as constituting a strong source of error in multiple choice tests like the TACL (Struening and Guttentag, 1975), is considered a component of the impulsive behavior sometimes exhibited by the 3 and 4-year-old age group (Ratusnik and Koenigskecht, 1975). In addition, the consideration that contribution of error was made by the motivation of the subject at the time of testing is also feasible (Ratusnik and Koenigskecht, 1975).

Another factor to consider as a source of error attributable to subject performance is variation due to differences of language ability exhibited by the black, low socioeconomic preschool group (Adler, 1971; Baratz, 1969; Cazden, 1966). However, since the literature related to this area is sparse, and studies which are available are related more to

production than comprehension tasks (with which the TACL is concerned), this factor has only been briefly considered. Variation in performance of formal test behavior has been noted within this group (Sigel and Perry, 1968). It is difficult to state a conclusion regarding its effect on the contribution of error within a standardized assessment device due to conflict in the literature on the issue of low socio-economic children's language abilities. Some authors have attributed variation in performance of this group to a "psycholinguistic diversity" that exists among low socioeconomic, black children (Sigel and Perry, 1968; Ratusnik and Koenigskecht, 1975; Ervin-Tripp, 1971). Others have pointed out that we do not know enough about the language of this group to make statements of any kind regarding it (Somervill, 1974). While the literature concerning the language ability of low socioeconomic children is conflicting, language differences have a part in the scores received on the TACL.

The conclusion which is feasible is that there is not sufficient data to indicate that standardized tests measuring language comprehension such as the TACL are either valid or reliable in assessing the language performance of low socioeconomic, black children. Language assessment devices which are standardized for this group are greatly needed.

Content of Instrument

Lastly, in the consideration of sources of error in measurement, it is necessary to regard sources in the content of the instrument. Research has reported that content error typically constitutes the majority of measurement error (Kerlinger, 1964). For this reason, it is the most

difficult source to study. The sources of error within the TACL which may have contributed to the results of this study are those individual test items which interfere with language or dialect differences in the black, low socioeconomic child, or those items which were missed by the greatest percentage of the entire group tested. Another source of variation may have been presented in the line drawing stimuli themselves. Some drawings with a contrast or decoy are colored darker or made to appear more obvious than the correct referent, which could account for confusion in responding. Some stimuli words to be presented by the examiner are sometimes as closely related to the correct referent as they are to the contrast or decoy presented. An area of research could be to examine each item for variability.

In conclusion, the results of this research indicate that the Test for Auditory Comprehension of Language has only a moderate level of internal consistency for use with low socioeconomic, black children, a level which is less than desirable for satisfactory measurement of language comprehension in this group. These statistical results showed a very limited critical range of test scores and narrow standard deviations for each age group, contributing to resulting correlations of .59 for 3-year-olds and .66 for 4-year-olds. Raw scores fell between one and two standard deviations below those reported in the normative data established for the TACL. Data reported specifically for low socioeconomic, black children by Musselwhite (1975) and results of this study show black low socioeconomic children have different scores than reported in the norms of the TACL.

Furthermore, possible sources of error which may have contributed to the results of this study have been discussed. They indicate that there is no single explanation for not obtaining a higher level of internal consistency on the TACL as it is used with black, low socio-economic children, and because of this lack of explanation there is a need for further research into the structure of the test. An important area is an item analysis of the TACL using data from this study. It is hoped that as a result of this research, there will be a greater concern for construction of tests which reliably measure language comprehension in the low socioeconomic, black population by providing normative data pertinent to that group and with data regarding the internal consistency. A concern for issues such as these is pertinent to language assessment and intervention in the black, low socioeconomic preschool children.

Chapter V

SUMMARY

Research has demonstrated the extreme importance of constructing measurement instruments which are relatively free from error. It has also shown that to the degree to which instruments possess this quality they demonstrate internal consistency, or are reliable in their measurement of performance. This study investigated the internal consistency of the Test for Auditory Comprehension of Language (TACL) (Carrow, 1973) as it is used in the diagnostic assessment of the language comprehension of black, low socioeconomic preschool children. Prior to this study, no research has been conducted into the internal structure of this test other than normative data collected in its construction.

Sixty black, Headstart children served as subjects. They included 30 3-year-olds and 30 4-year-olds. Subjects were administered the TACL after satisfactorily performing on a language screening test. Each subject was given the test at his or her Headstart center and responses were recorded for each subject on test record forms. Social reinforcement was delivered systematically for appropriate attending behavior throughout the test administration. Resulting raw scores were collected and analyzed for the tests according to the alpha coefficient method, the general form for estimating reliability in standardized tests. Correlation coefficients were obtained for each age group.

Results of this study indicate that the Test for Auditory Comprehension of Language has only a moderate level of internal consistency for use with low socioeconomic, black preschool children, a level less than desirable for satisfactory language assessment. This level was reported statistically in the correlation coefficients of .59 for 3-year-olds and .66 for 4-year-olds, a limited critical range of test scores, and narrow standard deviations for both age groups included in this study. The possible reasons for this moderate level of internal consistency for the TACL have been discussed. The sources of error include variation within subject performance such as guessing, state of motivation, and language differences. Error within the content of the test as a source of error has also been proposed in the form of specific item drawings and stimulus words. Future research into item analysis of the TACL has been suggested as a relevant study.

The results of this research indicate the importance and need for continued, extensive study into the internal structure of tests as they are used with groups such as the low socioeconomic, black populations. On a wider perspective, this study has emphasized the need for assessment devices which measure language performance according to normative data established for the group, and data reflecting internal consistency within the structure of the tests.

APPENDIX

TEST FOR AUDITORY COMPREHENSION OF LANGUAGE (TACL)

SAMPLE RECORD FORM

CHILD _____ DATE _____

ITEM #	CHILD RESP.	CORRECT RESP.	ITEM #	CHILD RESP.	CORRECT RESP.	ITEM #	CHILD RESP.	CORRECT RESP.	ITEM #	CHILD RESP.	CORRECT RESP.
*1		2	53		3	2		2	52		2
3		1	55		1	4		3	54		2
5		2	57		1	*6		3	*56		2
7		1	59		2	8		1	58		1
9		3	*61		3	10		2	60		1
*11		1	63		1	12		3	62		3
13		1	65		2	14		2	64		1
15		1	67		3	*16		3	*66		1
17		1	69		2	18		2	68		1
19		3	*71		3	20		1	70		1
*21		1	73		2	22		3	72		2
23		3	75		3	24		2	74		3
25		3	77		3	*26		3	*76		1
27		1	79		2	28		2	78		1
29		1	*81		1	30		1	80		1
*31		3	83		2	32		3	82		3
33		1	85		2	34		2	84		3
35		3	87		1	*36		2	*86		2
37		1	89		1	38		1	88		3
39		2	*91		1	40		3	90		2
*41		1	93		3	42		2	92		1
43		2	95		1	44		1	94		2
45		3	97		3	*46		2	*96		3
47		1	99		2	48		2	98		1
*49		1	*101			50		1	100		2
51		3									

REFERENCES

Adler, S., Dialectal differences: Professional and clinical implications. J. Speech Hearing Dis., 36, 1, 1971.

Baratz, J., A bi-dialectal task for determining language proficiency in economically disadvantaged Negro children. Child Development, 40, 889-902, 1969.

Baratz, J., The application of dialect research in the context of the classroom - it ain't easy. Acta Symbolica, 11, 1970.

Baratz, J., Language and cognitive assessment of Negro children. Assumptions and research needs. ASHA, 11, 1969.

Baratz, Jr., Language in the economically disadvantaged child: A perspective. ASHA, 10, 143-145, 1968.

Braun, J., A correlational study of the Northwestern Syntax Screening Test and the Test for Auditory Comprehension of Language. Master's Thesis, Moorhead State College, Minnesota, 1975.

Bruck, M., and Tucker, R., Social class differences in the acquisition of school language. Merrill-Palmer Quarterly, 20, 3, 1974.

Campbell, D., and Stanley, J., Experimental and Quasi-Experimental Designs for Research. Chicago: Rand-McNally, Inc., 1963.

Carrow, E., Test for Auditory Comprehension of Language. Learning Concepts, Austin, Texas, 1973.

Carrow, E., and Lynch, J., Comparison of semantic versus syntactic comprehension in three groups of linguistically deviant children. Unpublished manuscript, 1975.

Carrow, E., Auditory comprehension of English by monolingual and bilingual pre-school children. J. Speech Hearing Res., 15, 2, 1972.

Carrow, Sister Mary Arthur, The development of auditory comprehension of language structure in children. J. Speech Hearing Dis., 33, 2, 1968.

Cazden, C., Subcultural differences in child language: An interdisciplinary review. Merrill-Palmer Quarterly, 12, 185-200, 1966.

Dickie, J., and Bagur, S., Considerations for the study of language in young low-income minority group children. Merrill-Palmer Quarterly, 18, 1, 1972.

Edwards, A., Statistical Methods. New York: Holt, Rinehart, and Winston, Inc., 1967.

Foreit, K., and Donaldson, P., Dialect, race and language proficiency. Another dead heat on the merry-go-round. Child Development, 42, 1572-1574, 1970.

Frankenberg, W., Camp, B., Natta, P., and Demersseman, J., Reliability and stability of the Denver Developmental Screening Test. Child Development, 42, 5, 1971.

Gerber, S., and Hertel, C., Language deficiency of disadvantaged children. J. Speech Hearing Res., 12, 2, 1969.

Guilford, J., Fundamental Statistics in Psychology and Education. New York: McGraw-Hill, 1965.

Jones, B., A study of oral language comprehension of black and white middle and lower class, pre-school children using standard English and black dialect in Houston, Texas, 1972. Doctoral Dissertation, University of Houston, Texas, 1972.

Kerlinger, F., Foundations of Behavioral Research. New York: Holt, Rinehart, and Winston, Inc., 1964.

Levy, B., and Cook, H., Dialect proficiency and auditory comprehension in standard and nonstandard black English. J. Speech Hearing Res., 16, 1973.

Mecham, M., Jex, J., and Jones, J., Utah Test of Language Development. Communication Research Associates, Salt Lake City, Utah, 1967.

Miner, L., and Shriner, T., Morphological structures in the language of disadvantaged and advantaged children. J. Speech Hearing Dis., 11, 605-610, 1968.

Monsees, E., and Berman, C., Speech and language screening in a summer Headstart program. J. Speech Hearing Dis., 33, 2, 1968.

Musselwhite, C., Normative data on the Carrow Test for Auditory Comprehension of Language for lower socioeconomic status children. Unpublished manuscript. Operation Headstart, Durham, North Carolina, 1975.

Newton, R., Planning for the language development of disadvantaged preschool children and youth. J. Negro Education, 34, 1965.

Osser, H., Wang, M., and Zaid, F., The young child's ability to imitate and comprehend speech: A comparison of two subcultural groups. Child Development, 40, 1969.

Ramer, A., and Rees, N., Selected aspects of English morphology in black American children of low socioeconomic background. J. Speech Hearing Res., 16, 4, 1973.

Raph, J., Language and speech deficits in culturally disadvantaged children: Implications for the speech clinician. J. Speech Hearing Dis., 32, 1967.

Ratusnik, D., and Koenigskecht, R., Influence of certain clinical variables on black preschoolers nonstandard phonological and grammatical performance. J. Comm. Dis., 8, 281-297, 1975.

Sigel, I., and Perry, C., Psycholinguistic diversity among culturally deprived children. Amer. J. Orthopsychiatry, 38, 122-126, 1968.

Somervill, M., Language of the disadvantaged: Toward resolution of conflict. J. Negro Education, 43, 284-301, 1974.

Struening, E., and Guttentag, M., Handbook of Evaluation Research., Vol. 1, California Sage Pub., 1975.

Thorndike, R., and Hagen, E., Measurement and Evaluation in Psychology and Education. Wiley and Sons, Inc., 1969.

Weiner, P., and Hoock, W., The standardization of tests: Criteria and criticism. J. Speech Hearing Res., 16, 1973.

Wisk, J., Educational Measurement. New York: Merrill Pub., 1973.

Williams, F., Reasoning with Statistics: Simplified Examples in Communications Research. New York: Holt, Rinehart, and Winston, Inc., 1968.

Williams, F., Psychological correlates of speech characteristics...on sounding "disadvantaged". J. Speech Hearing Res., 13, 4, 1970.